

К.Е. БУРНАЕВ, БелГТУ им. Шухова (Белгород, Россия)

МОДЕЛИРОВАНИЕ НЕСТАЦИОНАРНЫХ ПОЛЕЙ В НЕОДНОРОДНЫХ ОБЛАСТЯХ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

У статті розглядається можливість отримання наближеного рішення диференціальних рівнянь у частинних похідних за допомогою одношарових нейронних мереж із зворотними зв'язками. Представлений приклад застосування мереж вказаної структури для вирішення диференціального рівняння.

The possibility of solving partial derivative differential equations by using neural networks with specific structure is investigated in the article. The simple single-layer neural network ability to approximate the solution of PDE is proved. A simple example of described method application is given.

Постановка проблемы. Задана n -мерная область Ω произвольной формы и интервал времени $[0, t]$. На границе Γ области определены краевые условия $F(\vec{r}, t), \vec{r} \in \Gamma$. Дано дифференциальное уравнение в частных производных с переменными коэффициентами

$$\frac{\partial}{\partial t} F(\vec{r}, t) = \frac{\partial}{\partial r} \left(k(\vec{r}, t) \frac{\partial}{\partial r} F(\vec{r}, t) \right), \quad (1)$$

где \vec{r} – произвольный n -мерный вектор, $\vec{r} \in \Omega$, $k(\vec{r}, t)$ – известная гладкая функция, определенная на области Ω , удовлетворяющая условию

$$0 \leq c_1 \leq k(\vec{r}, t) \leq c_2. \quad (2)$$

Необходимо найти функцию $F(\vec{r}, t)$.

С точки зрения практики, решение такого дифференциального уравнения соответствует моделированию нестационарного поля в среде с неоднородными характеристиками – например, температурного поля в объекте с непостоянной теплопроводностью или плотностью. Уравнения этого класса часто встречаются при решении практических задач; хотя общие принципы решения этих задач исследованы, алгоритмы решения остаются крайне требовательными к вычислительным ресурсам.

Анализ литературы. Несмотря на то, что принципы решения таких дифференциальных уравнений хорошо изучены, сам процесс решения представляется сложным. В большинстве случаев задача сводится к сверхбольшой системе линейных алгебраических уравнений, решение которой дает приближенное значение искомой функции в некотором наборе точек – узлов дискретизации. Заметим, что распространенные методы решения таких

систем являются итеративными и предполагают последовательную, а не параллельную обработку данных; таким образом, для ускорения процесса решения системы требуется увеличивать производительность процессора [1].

С другой стороны, в настоящее время начали получать всё большее распространение многоядерные процессоры, позволяющие использовать явный параллелизм вычислений даже при наличии только одного компьютера. Кроме того, большая часть организаций, которым могут потребоваться решение подобных задач, обладает большими количеством относительно маломощных рабочих станций, объединенных локальными сетями. Таким образом, оптимальное использование существующих ресурсов возможно только при максимальном распараллеливании процессов вычислений, что во многих случаях легко достигается при использовании нейронных сетей.

На текущий момент существует три основных направления применения нейронных сетей при решении дифференциальных уравнений в частных производных с постоянными коэффициентами. Первое использует идеи метода конечных элементов, а именно – приближение точного решения дифференциального уравнения с помощью набора простых базисных функций. Такая форма решения может быть отображена на архитектуру нейронной сети прямого распространения, в которой активационные функции нейронов соответствуют используемым базисным функциям, а весовые коэффициенты – коэффициентам, используемым при аппроксимации. Недостатком этого метода является то, что количество базисных функций очень велико и быстро растет с увеличением количества узлов задачи; при увеличении размерности задачи увеличивается количество параметров, необходимых для задания базисных функций. Это приводит к тому, что метод плохо масштабируется и начинает проигрывать обычным численным методам при решении задач большой размерности [2].

Второе направление – решение полученной при применении метода конечных разностей системы линейных алгебраических уравнений обращением матрицы с помощью обучаемой нейронной сети с архитектурой Хопфилда. Как известно, обучение сети с такой архитектурой соответствует минимизации энергетической функции. Если полученную систему линейных алгебраических уравнений отобразить на нейронную сеть так, чтобы энергетическая функция соответствовала невязке решения системы линейных уравнений, в результате обучения нейронной сети будет получено приближенное решение системы. Недостатком этого метода является большое время обучения и то, что обученную нейронную сеть нельзя применить для решения другой системы уравнений [3, 4].

Третье направление – получение дифференцируемого аналитического решения с помощью нейронных сетей специфической структуры. В отличие от предыдущих двух методов, этот метод дает приближенное аналитическое, а не численное решение поставленной задачи, представленное в виде суперпозиции двух функций

$$U(x) = A(x) - F(x, N(x, p)), \quad (3)$$

где функция $A(x)$ удовлетворяет краевым условиям, а $F(x, y)$ обращается на границе области в ноль и частично задается нейронной сетью $N(x, p)$, являющейся обычным трехслойным перцептроном, обучаемом на наборе заранее известных точек решения. Эта методика требует, чтобы область, в которой производится поиск решения, была прямоугольной, что не всегда возможно. Кроме того, для обучения перцептрона требуется наличие решения задачи на более грубой сетке, что подразумевает использование какого-либо другого метода в связке с этим [5, 6].

Отметим, что все подходы, перечисленные выше, предполагают, что решается дифференциальное уравнение с постоянными коэффициентами. Это накладывает серьезные ограничения на круг решаемых задач.

Цель статьи: предложить способ решения дифференциальных уравнений в частных производных с переменными коэффициентами, который бы обладал возможностью простого распараллеливания вычислений и в то же время точностью, сравнимой с распространенными универсальными численными методами.

Предлагаемая методика. Для решения задачи используем ячеистую нейронную сеть. Ячеистые нейронные сети представляют собой большие массивы одинаковых динамических систем – ячеек, которые удовлетворяют следующим требованиям: 1) все взаимодействия ограничиваются некоторым конечным радиусом и 2) переменные, описывающие состояние нейронной сети, являются непрерывными [7, 8].

Выделим в области Ω N узлов дискретизации с координатами $\vec{r}_1, \dots, \vec{r}_N$, образующие n -мерную однородную прямоугольную сетку дискретизации с шагом h_i по i -й координате; разобьем временной промежуток на m отрезков длиной τ .

Структуру сети можно описать следующим образом: сеть состоит из m слоев, каждый из которых представляет собой n -мерную прямоугольную решетку нейронов, имеющих линейные активационные функции $f(x) = x$. Каждый нейрон с координатами (r_1, \dots, r_n) внутри слоя соответствует узлу дискретизации с координатами $(r_1 h_1, \dots, r_n h_n)$. Слой с номером i , $i \in (0, m-1)$ соответствует моменту времени $i\tau$. Каждый нейрон имеет $2n$ связей с непосредственными соседями внутри слоя, обратную связь и две связи с нейронами с тем же номером в предыдущем и последующем слоях (за исключением нейронов первого и последнего слоя, у которых одна из этих связей будет отсутствовать).

Введем дополнительные вспомогательные обозначения: номер нейрона в слое i с координатами \bar{x} внутри слоя – $L(x, i)$; номер нейрона, предшествующего или последующего i -му по j -й координате – $P(i, j)$ и $S(i, j)$ соответственно; выходное значение i -го нейрона обозначим как f_i ; множество номеров нейронов, соседствующих с нейроном i – $B(i)$.

Определим веса связей, которые будут использоваться в сети. Вес обратной связи:

$$w_{ii} = b - \frac{1-b}{k_i k_i}, \quad (4)$$

где $k_i = k(r_i, t)$ – значение коэффициента, характеризующего свойства среды в данной точке (например, величины теплопроводности при моделировании теплового поля), k_i – вспомогательная величина, определяемая выражением

$$k_i = k(r_i, t) = \frac{\tau}{h^2} \sum_{i=1}^n \left(\frac{2k_{L(P(\bar{r}, i), t)}}{k_{L(P(\bar{r}, i), t)} + k_{L(\bar{r}, t)}} + \frac{2k_{L(S(\bar{r}, i), t)}}{k_{L(S(\bar{r}, i), t)} + k_{L(\bar{r}, t)}} \right). \quad (5)$$

Вес связи между нейронами внутри слоя:

$$w_{ij} = \begin{cases} \frac{1-b}{k_j} \cdot \frac{2k_i}{k_i + k_j}, & i \in B(j), \\ 0, & i \notin B(j). \end{cases} \quad (6)$$

Вес связи между нейронами разных слоев:

$$w_{L(\bar{x}, i) L(\bar{x}, i+1)} = \frac{1-b}{k_{L(\bar{x}, i)} k_{L(\bar{x}, i)}}. \quad (7)$$

Заметим, что, в отличие от связей внутри слоя, межслойные связи являются односторонними.

Используем информацию о граничных условиях следующим образом: сделаем активационные функции нейронов, соответствующим узлам дискретизации, для которых выполняется условие $f(\bar{r}, t) = \mu(\bar{r}, t) = \text{const}$, константными – $f_i(x) = \mu_i = \text{const}$. Аналогично изменим активационные функции нейронов, соответствующих узлам, в которых задано граничное условие вида $f'(\bar{r}, t) = \mu(\bar{r}, t) = \text{const}$ на функции вида:

$$f_i(x) = k \left(x + \frac{v}{2n_i} \right), \quad (8)$$

и вес обратной связи:

$$w_{ii} = \frac{b}{\alpha}, \quad (9)$$

где $\alpha = \frac{2n}{|B(i)|}$ – коэффициент, зависящий от количества соседних узлов.

Заметим, что стабильное состояние сети удовлетворяет конечно-разностному представлению дифференциального уравнения

$$\begin{aligned} & \frac{f(x, t) - f(x, t + \tau)}{\tau} = \\ & = \frac{1}{h} \left(a(x + h, t) \frac{f(x + h, t) - f(x, t)}{h} - a(x, t) \frac{f(x, t) - f(x - h, t)}{h} \right), \end{aligned} \quad (10)$$

где $a(x, t)$ – разностный коэффициент, вычисляемый по формуле:

$$a(x, t) = \frac{2k(x + h, t)k(x, t)}{k(x + h, t) + k(x, t)}, \quad (11)$$

вследствие того, что равенство

$$\begin{aligned} f_{L(\bar{x}, t)} &= w_{L(\bar{x}, t)L(\bar{x}, t)} f_{L(\bar{x}, t)} + w_{L(\bar{x}, t-1)L(\bar{x}, t)} f_{L(\bar{x}, t-1)} + \\ &+ \sum_{i=1}^n \left(w_{L(P(\bar{x}, i), t)} f_{L(P(\bar{x}, i), t)} + w_{L(S(\bar{x}, i), t)} f_{L(S(\bar{x}, i), t)} \right), \end{aligned} \quad (12)$$

описывающее стабильное состояние нейронной сети, может быть преобразовано к виду:

$$\begin{aligned} f_{L(\bar{x}, t)} &= \left(b - \frac{1-b}{k_{L(\bar{x}, t)}k_{L(\bar{x}, t)}} \right) f_{L(\bar{x}, t)} + \frac{1-b}{k_{L(\bar{x}, t)}k_{L(\bar{x}, t)}} f_{L(\bar{x}, t-1)} + \\ &+ \frac{1-b}{k_{L(\bar{x}, t)}} \sum_{i=1}^n \left(\frac{2k_{L(P(\bar{x}, i), t)}}{k_{L(\bar{x}, t)} + k_{L(P(\bar{x}, i), t)}} f_{L(P(\bar{x}, i), t)} + \frac{2k_{L(S(\bar{x}, i), t)}}{k_{L(\bar{x}, t)} + k_{L(S(\bar{x}, i), t)}} f_{L(S(\bar{x}, i), t)} \right); \end{aligned} \quad (13)$$

$$\begin{aligned} f_{L(\bar{x}, t)} - f_{L(\bar{x}, t-1)} &= k_{L(\bar{x}, t)}k_{L(\bar{x}, t)} f_{L(\bar{x}, t)} + \\ &+ k_{L(\bar{x}, t)} \sum_{i=1}^n \left(\frac{2k_{L(P(\bar{x}, i), t)}}{k_{L(\bar{x}, t)} + k_{L(P(\bar{x}, i), t)}} f_{L(P(\bar{x}, i), t)} + \right. \\ &\left. + \frac{2k_{L(S(\bar{x}, i), t)}}{k_{L(\bar{x}, t)} + k_{L(S(\bar{x}, i), t)}} f_{L(S(\bar{x}, i), t)} \right); \end{aligned} \quad (14)$$

$$f_{L(\bar{x},t)} - f_{L(\bar{x},t-1)} = k_{L(\bar{x},t)} \left(\frac{\tau}{h^2} \right) \sum_{i=1}^n \left(\frac{2k_{L(\bar{x},t)}k_{L(P(\bar{x},i),t)}}{k_{L(\bar{x},t)} + k_{L(P(\bar{x},i),t)}} (f_{L(P(\bar{x},i),t)} - f_{L(\bar{x},t)}) + \right. \\ \left. + \frac{2k_{L(\bar{x},t)}k_{L(S(\bar{x},i),t)}}{k_{L(\bar{x},t)} + k_{L(S(\bar{x},i),t)}} (f_{L(S(\bar{x},i),t)} - f_{L(\bar{x},t)}) \right). \quad (15)$$

Граничные условия выполняются в силу того, что активационные функции нейронов, соответствующих граничным узлам, являются константными. Таким образом, стабильное состояние сети эквивалентно приближенному численному решению поставленной задачи.

Единственность стабильного состояния очевидным образом вытекает из приведенных выше преобразований. Покажем, что это стабильное состояние сети достижимо. Введем следующие вспомогательные обозначения: W – матрица весов сети, F – вектор выходных значений нейронов сети, D – вектор единичных смещений нейронов. Заметим, что в силу линейности активационных функций, $F_{t+1} = WF_t + D$.

Докажем, что сеть не может бесконечно осциллировать между несколькими состояниями с помощью доказательства от противного. Пусть существует набор попарно различных состояний сети F_1, \dots, F_n такой, что

$$F_2 = WF_1 + D, \dots, F_1 = WF_n + D. \quad (16)$$

Тогда должно выполняться следующее равенство:

$$F_1 = F_1 W^n + D \left(\sum_{i=1}^{n-1} W^i \right); \quad (17)$$

$$(E - W^n)F_1 = D \left(\sum_{i=1}^{n-1} W^i \right). \quad (18)$$

Заметим, что для любого n $W^n \neq E$. Решая полученное матричное уравнение относительно F , получаем

$$F_1 = \frac{D \left(\sum_{i=1}^{n-1} W^i \right)}{(1 - W) \sum_{i=1}^{n-1} W^i} = \frac{D}{1 - W}; \quad (19)$$

$$F_2 = F_1 W + D = \frac{D}{1 - W} W + D = \frac{D}{1 - W} = F_1, \quad (20)$$

что противоречит начальному предположению о том, что состояния F_i попарно неравны.

Покажем, что выходные значения нейронов сети ограничены. В общем случае, когда нейроны могут иметь и константные, и линейные активационные функции, выходные значения нейронов на каждом шаге можно определить с помощью выражения

$$F_{t+1} = F_t W L + C, \quad (21)$$

где C – вектор константных выходных значений, содержащий нулевые значения в элементах, соответствующих нейронам с линейными активационными функциями; L – квадратная матрица размером n , элементы которой определяются по следующему правилу:

$l_{ii} = 1$, если i -й нейрон имеет линейную активационную функцию,

$l_{ij} = 0$ во всех прочих случаях.

Заметим, что $CWL \equiv 0$. Тогда

$$F_t = F_0 (WL)^t + C. \quad (22)$$

Так как $\|W\| \leq 1$, то и $\|WL\| \leq 1$. Следовательно

$$\sum_{i=1}^N |f_i^{(t+1)}| \leq \sum_{i=1}^N |f_i^{(t)}| + \sum_{i=1}^N c_i \leq \sum_{i=1}^N f_0 + \sum_{i=1}^N c_i \leq c. \quad (23)$$

Очевидно, что никакой выходной сигнал не будет по модулю превышать этого постоянного значения.

Пример. Для иллюстрации рассмотренной методики используем уравнение распространения теплового импульса в составном стержне с кусочно-линейным коэффициентом теплопроводности.

$$\begin{aligned} \frac{\partial f}{\partial t} &= k^2 \frac{\partial^2 f}{\partial x^2}, \\ \left\{ \begin{aligned} f(0,0) &= f_0; \\ f(0,t) &= 0; \\ f(l,t) &= 0; \\ f(x,0) &= 0; \end{aligned} \right. \\ f_0 &= 10. \end{aligned}$$

Будем считать, что стержень состоит из частей длины $0.l=1$ с чередующимися коэффициентами теплопроводности $k_1=10$ и $k_2=20$.

Результат решения задачи с применением нейронных сетей приведен на рисунке. Решение, полученной с помощью нейронной сети, соответствующей равномерной сетке дискретизации, состоящей из $100 = 10 \cdot 10$ узлов, отклоняется от аналитического на величины порядка 10^{-6} , сравнимые с величиной, используемой в критерии завершения работы сети.

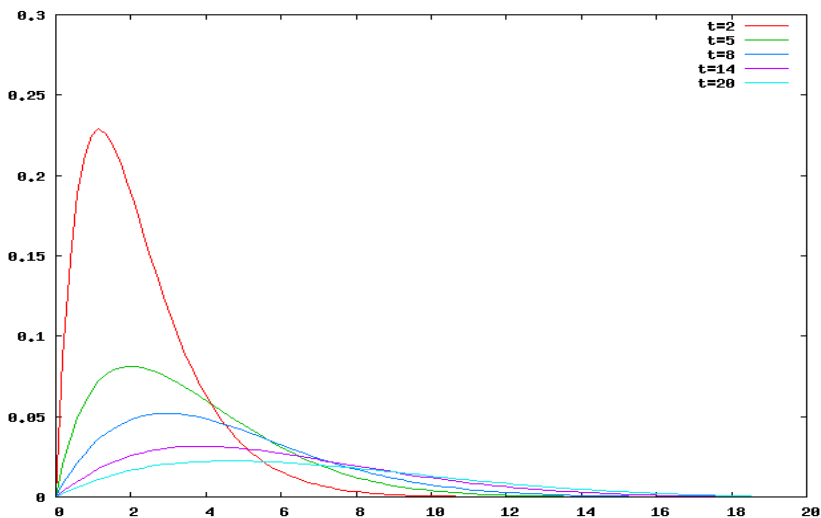


Рис. Общий вид решения примера

Выводы. Предложенная нейронная сеть позволяет получать приближенное численное решение параболических дифференциальных уравнений в частных производных с переменными коэффициентами. В силу небольшого радиуса взаимодействия между нейронами в сети, предложенной структуры, вычисления, использующие её, могут быть эффективно распараллелены.

Список литературы: 1. Арфкен Г. Математические методы в физике. – М: Атомиздат, 1969. – 712 с. 2. Lee H., Kang I. Neural algorithms for solving differential equations // Journal of Computational Physics. – 1990. – V. 91. – P.110–117. 3. Wang L., Mendel J.M. Structured trainable networks for matrix algebra // IEEE Int. Joint Conference on Neural Networks. – 1990. – V. 2. – P. 125–128. 4. Gupta A., Karypis G., Kumar V. Highly Scalable Parallel Algorithms for Sparse Matrix Factorizations // IEEE Transactions on Parallel and Distributed Systems. – 1995. – V. 8. – № 5. – P. 502–520. 5. Lagaris I.E., Likas A., Fotiadis D.I. Artificial Neural Networks for Solving Ordinary and Partial Differential Equations // IEEE Transactions on Neural Networks. – 1998. – V. 9. – № 5. – P. 422–448. 6. Lagaris I.E., Likas A., Fotiadis D.I. Artificial Neural Networks for Solving Ordinary and Partial Differential Equations // IEEE Trans. on Neural Networks. – 1998. – V. 4. – P. 987–1000. 7. Rekeczky C., Szatmári I., Földesy P., Roska T. Analogic Cellular PDE Machines. – Computer and Automation Research Institute. Hungarian Academy of Sciences, Budapest, Hungary. Tech. Report, 2004. 8. Fortuna L., Arena P., Balya D., Zarándy A. Cellular Neural Networks: A Paradigm for Nonlinear Spatio-Temporal Processing // Circuits and Systems Magazine, IEEE. – 2001. – V. 1. – № 4. – P. 6–21.

Поступила в редакцию 11.10.2005